DOCUMENT RESUME

ED 440 515 EC 307 805

AUTHOR	Marion, Scott F.; Sheinker, Alan
TITLE	Issues and Consequences for State-Level Minimum Competency Testing Programs. State Assessment Series. Wyoming Report 1.
INSTITUTION	National Association of State Directors of Special Education, Alexandria, VA.; National Center on Educational Outcomes, Minneapolis, MN.; Council of Chief State School Officers, Washington, DC.
SPONS AGENCY	Special Education Programs (ED/OSERS), Washington, DC.; Wyoming State Dept. of Education, Cheyenne.
PUB DATE	1999-01-00
NOTE	36p.
AVAILABLE FROM	NCEO Publications Office, 350 Elliott Hall, 75 E. River Road, Minneapolis, MN 55455; Tel: 612-624-8561; Fax: 612-624-0879; Web site: http://www.coled.umn.edu/nceo ($10).
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Standards; *Disabilities; Educational Assessment; Elementary Secondary Education; Instructional Effectiveness; *Minimum Competency Testing; *State Programs; *Student Evaluation; *Theory Practice Relationship
IDENTIFIERS	Wyoming

ABSTRACT

This report reviews the current status, empirical findings, theoretical issues, and practical considerations related to state-level minimum competency testing programs. It finds that, although two-thirds of current testing programs now use direct writing prompts to assess writing achievement, essentially all programs rely on multiple choice tests to measure knowledge in the other subject areas. It also concludes that empirical evidence regarding the effectiveness of minimum competency testing programs is mixed. It reports improved achievement in basic reading and mathematics skills, especially when curriculum focuses on the same basic skill items found on the tests. However, the report also finds evidence of unintended negative effects of minimum competency testing programs, including lack of transfer to higher order skills, increased dropout rates (especially for minority and low achieving students), a narrowing of the curriculum to test content, corruptibility of high stakes tests, and testing time as time taken from teaching. Overall, the report finds a conflict between minimum competency and standards-based assessment systems since competency testing essentially contradicts current mandates for having students learn rigorous content standards. It recommends against mandating a state-level minimum competency program. (Contains 28 references.) (DB)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 440 515

STATE ASSESSMENT SERIES
Wyoming Report 1

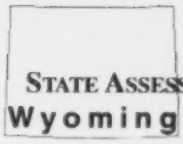
Issues and Consequences for State-Level Minimum Competency Testing Programs

EC 307 805

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE



STATE ASSESSMENT SERIES
Wyoming Report 1

Issues and Consequences for State-Level Minimum Competency Testing Programs

Prepared By:

Scott F. Marion • Alan Sheinker
Wyoming Department of Education

January, 1999

In response to a 1997 Wyoming legislative mandate, the Wyoming Department of Education was charged with investigating the feasibility and appropriateness of implementing a high-stakes minimum competency testing program. The following paper presents the results of this inquiry. In part, as a result of the findings and recommendations in this paper, the Wymong Legislature dropped the minimum competency proposal from legislation authorizing the Wyoming Comprehensive Assessment System (WyCAS). WyCAS is an assessment system that contains both standards-based (tied to the Wyoming Content and Performance Standards) and norm-referenced assessments with a major focus on providing information on school improvement.

Permission is granted to copy this document without charge. Additional print copies may be purchased for $10.00 from:

NATIONAL CENTER ON EDUCATIONAL OUTCOMES
University of Minnesota • 350 Elliott Hall
75 East River Road • Minneapolis, MN 55455
612.624.8561 • Fax 612.624.0879
http://www.coled.umn.edu/NCEO

This document is available in alternative formats upon request.

Executive Summary

This report presents a thorough review of the current status, empirical findings, theoretical issues, and practical considerations related to state-level minimum competency testing programs. There is not a clear consensus about the purposes, consequences, and performance levels associated with multiple choice tests among states currently using them. The rhetoric describing these programs would have one believe that these tests no longer measure minimum levels of academic content; however, the evidence presented in this paper contradicts much of the rhetoric. Though two-thirds of current testing programs are now using direct writing prompts to assess students' writing achievement, essentially all of these programs rely on multiple choice tests to measure knowledge in the other subject areas tested. These tests are undoubtedly better than their predecessors in the 1970s, but many of them still appear to be a long way from assessing the types of knowledge and skills required for work in the 21st century.

The empirical evidence regarding the effectiveness of minimum competency testing programs appears mixed, but this appearance depends on the perspective from which one evaluates these studies. Several studies (e.g., Frederiksen, 1994, Winfield, 1990) reported that student achievement improves when minimum competency testing programs are implemented. However, the measure of achievement in these cases is usually the same type of items as found on minimum competency tests. If the curriculum is focused on teaching basic skills, students will most likely improve in this area, but the importance of this finding is questionable. The more important claim would address whether focusing on basic skills leads to an improvement in the types of high-level work place skills deemed important by business and labor groups.

Many of the studies reviewed in this paper provide clear evidence of negative effects of minimum competency testing programs. For example, Griffin and Heidorn (1996) showed that minimum competency tests do not help those they are most intended to help—students at the lowest end of the achievement distribution. They have a negative effect on students with some promise. In addition to leading to an increase in dropout rates, one of the most deleterious effects of minimum competency tests is their influence on curriculum at the school and classroom level. There have been several studies focusing on the effects of minimum competency testing on curriculum, teaching, and learning. Most of these studies have been critical of minimum competency tests because of their negative effects on curriculum and instruction.

Besides the tremendous costs involved, a fundamental problem with implementing both minimum competency and standards-based assessment systems at the same time is the resulting instructional confusion. Competency testing essentially contradicts current mandates for having students learn rigorous content standards. With dual systems and two different sets of expectations, teachers and students will be unsure of where they should aim. History informs us that the

testing program with the higher stakes will exert more of a driving force on curriculum and instruction.

From a practical standpoint, if the competency tests are aligned with the state standards then they will be unnecessary. The scoring system on the standards-based assessments will be designed to provide information about students scoring at levels all along the achievement continuum. Therefore, an additional testing program, besides creating confusion, will draw important resources away from a standards-based assessment designed to promote high levels of academic achievement for *all* students. A well-designed standards-based assessment system should be able to assess all students at their current achievement levels but should also be able to provide a clear set of demanding instructional goals for educators and students. We strongly recommend against mandating a state-level minimum competency program. The potential benefits are unclear or lacking, and there is the serious possibility of doing more harm than good.

Assessment plays a critical part within the context of current educational reforms. There are many benefits to be derived from setting standards for what students are to know and be able to do, and from assessing their performance against these standards. It is within this positive context of high expectations that states are most likely to move forward in their educational reform agenda.

Table of Contents

Executive Summary i
Overview 1
History and Policy Context of Minimum Competency Testing 1
Theoretical Assumptions of Minimum Competency Programs 3
Workplace Readiness 3
Beliefs about School Reform 5
Beliefs About Learning 6
Current Status of Minimum Competency Programs in the United States 8
Classifying Minimum Competency Programs 8
Characteristics of Current Programs 11
Types of Tests and Items 12
Consequences Associated with Current State Minimum Competency Programs 13
Summary of Current Programs 13
Evaluations of Minimum Competency Testing Programs 14
Introduction and Evaluation Criteria 14
Intended Effects 15
Increase in reading achievement 15
Increase in basic mathematical skills 15
Unintended Effects 16
Lack of transfer to higher-order skills 16
Increasing dropout rates 16
Narrowing the curriculum 18
Corruptibility of high stakes tests 19
Time away from teaching 19
Practical Issues 19
Financial commitment 19
Two Assessment Systems 21
Discussion 22
References 27

Overview

We are frequently bombarded with media stories of students who make it all the way through high school without being able to read, or that the cashier at a local convenience store is unable to make change when the computer loses power. Policymakers and citizens feel compelled to rectify the ills in the educational system so that these examples become rare exceptions. Minimum competency tests are often used as a policy tool to require that students meet some basic level of achievement, usually in reading, writing, and computation, with the intention that the use of these tests will lead to the elimination of these educational problems. The Fifty-Fourth Legislature of the State of Wyoming, 1997 Special Session, requested the Wyoming Department of Education, in Enrolled Act No. 2 of the House of Representatives, Section 203d, to "...consider and ... report to the legislature no later than December 1, 1998 on the cost and feasibility of minimum student competency tests after the fifth, seventh, and eleventh grade levels (State of Wyoming Legislature, 1997, p. 23)." Minimum competency tests have a long history in educational testing, thus there is a fairly extensive literature concerning this issue. The purpose of this report, in fulfilling the legislative requirement, is to review this literature focusing particularly on evaluations of minimum competency testing programs.

Conflicts could surface as a result of the state implementing two distinct testing programs. These conflicts are presented in greater detail later in the paper, but it is important to foreshadow this discussion here. The potential advantages of a minimum competency program cannot be evaluated in isolation. The State is required by law to implement a standards-based comprehensive assessment system, therefore any potential effects of minimum competency testing must be evaluated in terms of the benefits they can provide over and above the proposed Wyoming Comprehensive Assessment System.

Following this review, some of the practical considerations are discussed. These include costs associated with testing and remediation, the logistics of implementing appropriate remedial programs, links with standards and standards-based assessments associated with implementing a minimum competency testing program in Wyoming. In the last section of this report, conclusions and recommendations based on the synthesis of the literature review and practical considerations are presented.

History and Policy Context of Minimum Competency Testing

Over one hundred years ago, the British instituted the "Revised Code," a statute that took minimum competency testing to an extreme. Schools received funds for only those students meeting minimum attendance requirements and passing competency exams in reading, writing,

and arithmetic. Schools had to forfeit approximately twenty-five percent of their per-capita allotment for each student failing these exams (Tuman, 1979).

Historical lessons are often ignored, in spite of the fact that there is much to learn. This British example is a case in point. The Revised Code was enacted in response to specific economic and political pressures. In England, state aid to education was undergoing a period of rapid growth and taxpayers and legislators were demanding accountability for their increasing educational expenditures. The Code remained in effect for more than 30 years. It lost its statutory power because of the damaging effect it was having on teachers. Teaching became simply an effort to get students through these minimum competency hoops, without providing any noticeable increase in student learning (Tuman, 1979).

Jumping forward to the post-Sputnik era when many educational reforms—initially targeted to science and mathematics education—were instituted in the United States, minimum competency programs became popular in this country. The 1965 Elementary and Secondary Education Act (ESEA) included provisions for the evaluation of programs funded under this act. These evaluations usually included standardized tests to measure student achievement. Because continued funding depended on successful evaluations, there was increased pressure to "teach the test" (Frederiksen, 1994).

Using standardized tests to inform policymakers about educational programs spread beyond those funded under ESEA. Many states enacted legislation that required schools to administer minimum competency achievement tests. Two main reasons were often cited to support this legislation: "(a) to determine the level of basic skills at various grade levels, and (b) to provide a basis for remediation in schools where it is needed (Frederiksen, 1994, p.3)." Problems with these early minimum competency testing programs arose because there was not an agreed upon definition of minimum competency; as a result, policymakers defined these requisite skills within each jurisdiction (Winfield, 1990).

In spite of some of these concerns, minimum competency testing programs proliferated in the United States. Through the 1970s and early 1980s more than 40 states had instituted some form of minimum competency testing due to increasing demands for public accountability of schools (Frederiksen, 1994; Winfield, 1990). The 1983 publication of *Nation at Risk* (National Commission on Excellence in Education, 1983) undoubtedly raised concerns that United States' students were not being sufficiently educated to help this country compete in the global economy. This publication solidified concerns that low academic standards and policies of social promotion (i.e., promoting students from one grade to the next so they can remain with their same-age peers) led to degradation of job skills. Reports that many high school graduates could not read or perform simple mathematics operations fueled the minimum competency fire (Jaeger, 1982). Minimum competency testing policies were originally intended to add meaning to a high school

diploma; i.e., students had to demonstrate at least minimum levels of knowledge and skills if they were to graduate or at least move on to the next grade level. For graduation, it was assumed that these minimum levels would translate into successful job performance.

Another reason for the increased use of large-scale standardized achievement tests was that as a result of the intense focus on education, state legislatures were appropriating relatively more funds for education. These and other testing programs were instituted to provide the public and policymakers with a method for holding schools accountable for these increased expenditures. The logic beyond these arguments is that if schools are held accountable for their students' test results, they are bound to change so that increased student learning will result.

Theoretical Assumptions of Minimum Competency Programs

There are certain crucial assumptions operating when considering a minimum competency testing program. These fall into three main categories of assumptions: workplace readiness, school reform, and learning theory.

Workplace Readiness

Much of the push for the use of basic skills minimum competency testing programs has been tied to issues of work force readiness (e.g., Jaeger, 1982). That being the case, it is important to define this elusive construct. What is meant by workplace readiness? Is it a projection for the near future (e.g., 2001) or the more distant future (e.g., 2020)? What sector of the work force is being targeted? Are students to be certified simply for low-wage service sector jobs or for jobs that require a high level of technical skills, but not as much as a postsecondary degree?

Fortunately, the Secretary of the U.S. Department of Labor undertook defining the skills and knowledge U.S. student will need to be successful in the work force (e.g., The Secretary's Commission on Achieving Necessary Skills [SCANS], 1991; SCANS, 1992). The Commission defined the types of knowledge and skills workers would require for a "high-performance economy characterized by high-skill, high-wage employment (SCANS, 1992, p. xiii)."

SCANS' initial report (1991), *What Work Requires of Schools*, outlined its conception of a high-performance workplace and its expectations for the types of skills and knowledge workers will need to succeed in the economic times of the near future. According to SCANS, workers will be expected to have a solid foundation in "basic" literacy and computational skills, but they also will need to have thinking skills so they can put this knowledge to work. Further, they will

require the types of personal qualities that enable people to become effective workers. SCANS was unwilling to assume that merely because students possessed these skills, they would become effective workers. They indicated that "high-performance workplaces also require other competencies: the ability to manage resources, to work amicably and productively with others, and to acquire and use information, to master complex systems, and to work with a variety of technologies" (p. xiii). These foundational skills and competencies are outlined in Table 1.

According to SCANS, there is little doubt that the high-performing workplace of the near future will require considerably more than minimum skills. SCANS took the next step and considered the implications for the educational systems and suggested how schools should function in order to meet the demands of the workplace.

Table 1. Workplace Know-How

Workplace Competencies: Effective workers can productively use:

* **Resources**—They know how to allocate time, money, materials, space, and staff.
* **Interpersonal skills**—They can work on teams, teach others, serve customers, lead, negotiate, and work well with people from culturally diverse backgrounds.
* **Information**—They can acquire and evaluate data, organize and maintain files, interpret and communicate, and use computers to process information.
* **Systems**—They understand social, organizational, and technological systems; they can monitor and correct performance; and they can design or improve systems.
* **Technology**—They can select equipment and tools, apply technology to specific tasks, and maintain and troubleshoot equipment.

Foundational Skills: Competent workers in the high-performance workplace need:

* **Basic Skills**—Reading, writing, arithmetic and mathematics, speaking and listening.
* **Thinking Skills**—The ability to learn, to reason, to think creatively, to make decisions, and to solve problems.
* **Personal Qualities**—Individual responsibility, self-esteem and self-management, sociability, and integrity.

(From SCANS, 1992, p.6).

The commission offered the following recommendations:

- Teaching should be offered in context, that is, students should learn content while solving realistic problems. Learning in order to know should not be separated from learning in order to do.
- Improving the match between what work requires and what students are taught requires changing how instruction is delivered and how students learn.
- High performance requires a new system of school administration and assessment.
- The entire community must be involved (SCANS, 1992, p. 12).

These suggestions are not new. Progressive educators have been advocating for this type of education for many years. Hearing a similar cry from business and governmental communities is welcome news. Approaching teaching and learning in this way essentially prohibits the use of low-level minimum competency tests for several reasons. First, the type of learning communities called for by SCANS are based on up-to-date learning theories. This type of learning does not call for complete mastery of basic skills prior to moving onto more complex content. Second, calling for educational systems to focus on the teaching and learning of complex subject matter would be contradicted by simultaneous demands for minimum competency tests of discrete skills. Past history can provide an important lesson here. When high stakes (e.g., graduation, promotion) are attached to minimum competency tests, these tests receive an inordinate amount of teachers' and students' attention. Therefore, any laudable goals for improved teaching and learning will get pushed aside in order to make sure students succeed on the minimum competency test.

Beliefs about School Reform

Proponents of competency testing programs argue that use of such tests creates incentives for low-performing schools and students to improve their performance. This approach to reforming instruction and learning in our nation's schools is based on the premise that if students and teachers are aware of the specific content and performance standards and are held accountable through the use of a test or series of assessments, they will be motivated to meet these standards. This feature of having accountability tests influence or "drive" instruction is commonly referred to as Measurement-Driven Instruction (MDI). Minimum competency testing is merely a specific instance of measurement-driven instruction.

Popham (1987) coined the term Measurement-Driven Instruction (MDI) and viewed MDI as a way to help improve education in a fairly cost-effective way. He described conditions that

should lead to an effective MDI program and stated that if the following elements are present, an MDI program will likely have a beneficial effect : (a) use of high-stakes criterion-referenced tests to clearly specify the skills and knowledge required of students, (b) the tests and objectives should represent meaningful content, (c) a limited number of skills and objectives should be assessed, (d) teachers should be able to use the tested objectives to plan instruction, and (e) support for educators needs to be provided so that the tested knowledge and skills are able to be taught. The motivating factor of the high stakes test is responsible for driving the instruction toward the objectives and criterion defined by the test. Further, Popham (1987) advocated giving teachers prototype test items and urging them to teach all students how to pass these items.

There appears to be evidence that this strategy works, in the sense that it produces a rise in test scores. The Resnicks (1992) argued that whether we like it or not, what is taught and what is tested are intimately related. They suggested that every test used for accountability can affect the curriculum and offered the following principles to serve as guidelines for accountability assessments: "(1) you get what you assess... [and] (2) you do not get what you do not assess... (Resnick & Resnick, 1992, p. 59)."

Airasian (1988) noted that while the conditions set forth by Popham (1987) are important, they are incomplete, difficult to attain, and cannot regularly explain the variability among different MDI programs. High stakes, in and of themselves, are insufficient to drive instruction, rather it is the interaction of stakes, standards, and content that drive instruction. The greatest impact on instruction will occur when stakes and standards are both high (Airasian, 1988), as is the plan for some of the newer standards-based reforms. In addition, the more course-specific the test content, the more likely it is to drive instruction (Airasian, 1988).

While Airasian's (1988) argument about the interaction of standards, tests, and content is important, there is little debate about whether high stakes can drive instruction; rather, the controversy is concerned with whether they should. For instance Shepard (1991) and other critics (Haney & Madaus, 1989) have pointed out that efforts to improve the performance on particular tests seem to drive out most other educational concerns by progressively restricting the curriculum to the very narrow range of objectives tested. Most of these critics have been especially concerned with the effects of using traditional pencil and paper tests to test for minimum competency and to subsequently drive instruction.

Beliefs About Learning

In the social sciences there are often many theories or paradigms available to explain events and related phenomena. The explanations that people find compelling are shaped, to some extent, by the quality of supporting empirical evidence, but only within the context of specific

theoretical premises. For example, it is important to have at least an implicit theory (an explicit theoretical framework would be better) about certain phenomena in order to make sense of observations.

When policymakers and testing experts make decisions about certain assessment programs, their theoretical frameworks usually are not explicit. This is rarely a covert attempt to hide specific intentions; rather, most people are not aware of the theories that shape their observations. Nevertheless, decision-makers should be responsible for making their own learning theories explicit. More importantly, they should be willing to examine their theories in light of more recent scientific information. An examination of some of the learning theories and their relationship to testing and learning are presented in following paragraphs.

Most stakeholders agree that educational testing should, in some way, relate to improved opportunities for student learning. Therefore, if improved student learning is a goal, assumptions by which this learning is expected to occur need to be examined. Shepard's (1991) investigation of the implicit learning theories of a representative sample of district-level measurement specialists provides important insight into this question. Shepard found that approximately half of the measurement specialists adhered to a fairly strict criterion-referenced view of testing. This approach to testing is closely linked to behaviorist learning theory. Behaviorism expects learning in a given domain to be the sequential accumulation of requisite skills, therefore testing should occur at each specific learning step. While this may appear logical, there are many flaws.

Mislevy (1996) and others (e.g., Glaser & Silver, 1994; Shepard, 1991) have argued that recent advances in cognitive psychology—particularly in the areas of mental models, the nature of expertise, and situated learning—have implications for testing and test theory development. These new developments in human learning theory challenge the factory model of education and "the adequacy of the 'one size fits all' presumption of standard assessment" (Mislevy, 1996, p. 12). While many traditional notions of learning, where all students receive the same curriculum in the same fashion, are slowly falling by the wayside, large-scale, formal assessments are still designed and scored as if all students learned and progressed along the same instructional path.

Misunderstanding the nature of human learning has contributed to the resilience of certain assumptions about the relationship between basic skills and complex thinking. For example, many still believe that "basic" skills serve as the gateway to more complex understandings (or the "facts-before-thinking model of learning" according to Shepard, 1991). Certain basic skills are essential (e.g., simple computational skills), however current information from cognitive psychology indicates that students do not necessarily need to possess all of the essential basic skills before moving on to more complex content. Encountering interesting and complex problems

can lead students to learn "basic" material on the way to solving these problems. Use of a minimum competency test could reify the expectation that knowledge of basic skills must always precede learning more complex material.

Another troubling assumption underlying the use of competency tests relates to the expected unidimensionalilty of these domains (e.g., subtraction, spelling). This follows from the assumption of sequential mastery where the domain, as represented on the test, can be ordered along a difficulty (complexity) continuum. Items on minimum competency tests are generally targeted for a certain level along this slope. Holding fast to these assumptions may actually misrepresent student knowledge to the extent that students with "irregular" profiles are not given appropriate credit. Knowledge from cognitive psychology informs us that while there might be a hierarchical ordering to certain domains, most students do not have to acquire knowledge according to a perfectly ordered sequence.

Current Status of Minimum Competency Programs in the United States

In spite of their checkered history, competency testing programs in the United States are still alive and well. Most of these programs appear to be testing content at a slightly higher cognitive level than earlier versions of minimum competency testing. Table 2 presents an overview of current competency programs. These data were drawn from the Council of Chief State School Officers recent survey of state testing programs (CCSSO, 1996).

Classifying Minimum Competency Programs

Many states do not refer to their minimum competency testing programs as "minimum competency tests," therefore the following procedure was designed to identify and examine current state-level minimum competency testing programs. Any testing programs with the words "competency," "proficiency," "skills," or similar words or phrases were identified as potential minimum competency programs. In order to categorize assessments as competency tests, several criteria had to be met.

- These tests had to have individually reported scores with all respondents answering at least a common set of items (or all the same items).
- These tests had to be composed predominately of selection-type (e.g., multiple choice) or fill-in-the-blank questions.

Table 2. Characteristics of State Competency Testing Programs

State	Testing Program	Subject Areas	Grade Level	Grad. / Promote / Retest?	Type of Test	Item Type	Calc. Use	Performance Levels (#)
AL	Basic Competency Test	LA, M, R	9	Y / N / N	CRT	MC	N	2-Pass; Fail
AL	Exit Exam	LA, M, R	11, 12	Y / N / Y	CRT	MC	N	2-Pass; Fail
FL	H.S. Competency Test	CO, M	12	Y / N / Y	CRT	MC	Y	?
GA	GA H.S. Graduation Tests	LA, M, S, SS, W	11 > 12	Y / N /Y	CRT, WP	MC, SA, EA	Y	2-Pass; Fail
HI	Test of Essential Competencies	Basic Skills	10 > 11 > 12	Y / N / Y	CRT	MC	N	2-Pass; Fail
LA	Criterion - Referenced Test	LA, M	3, 5, 7	N / Y / N	CRT	MC	N	2-Pass; Fail
LA	LA Graduation Exit Examination	LA, M, S,SS, W	10 > 11	Y / Y / Y	CRT, WP	MC, ER	N	2-Pass; Fail
MD	MD Functional Testing Program	CI, R, W	7>8 >9 > 10>11>12	Y / N / Y	CRT, WP	MC, ER	N	3 - Excellent; Satisfactory; Fail
MS	Functional Literacy Examination	M, R, W-CO	11	Y / N / Y	CRT, WP	MC, ER, ESW	N	3 - Max. Acceptable, Min. Acceptable, Not Acceptable
NC	Competency Testing	LA, M, R	8, 10	Y / N / Y	CRT, NRT	MC	Y	2-Pass; Fail
NJ	Early Warning Test	M, R, W	8	N / N / N	CRT, WP	MC, SA, ER	Y	2-Pass; Fail
NJ	Grade 11 H.S. Proficiency Test	M, R, W	11 > 12	Y / N / Y	CRT, WP	MC, SA, ER	Y	3 - Clearly Competent; Min. Comp; Not Comp.
NM	NM H.S. Competency Examination Program	LA, M, R, S, SS, W	10 > 11 > 12	Y / N / Y	CRT, WP, PA	MC, MC-SE, SA, ER	Y	2-Pass; Fail
NV	H.S. Proficiency Exam. Program	M, R, W	11 > 12	Y / N / Y	NRT, WP	MC, ER	N	2-Pass; Fail
NY	Preliminary Competency Test	R, W	8 , 9	N / N / N	CRT, WP	FB/C, ER	N	2 - Below stds.; meets stds.
NY	Regents Competency Tests	M, R, S, SS, W	9-M&S, 10-SS, 11-R,W, SS	Y / N / Y	CRT, WP	MC, FB/C, ER	N	2 - Below stds.; meets stds.

Note: This table is a summary of information presented in the 1996 Council of Chief State School Officers (CCSSO) Annual Survey of State Student Assessment Programs. The information in this table was compiled from the following parts of the report: Part 3.02, pp. 119-130; Part 3.10, pp. 159-162; Part 3.11, pp. 163-167; Part 3.12, pp. 169-170; Part 3.13, pp. 171-175; Part 3.14A, pp. 176-180; Part 316, pp. 186-188; Part 3.20, pp. 200-204; Part 3.21, pp. 205-207.

Table 2. Characteristics of State Competency Testing Programs (continued)

State	Testing Program	Subject Areas	Grade Level	HS Grad. / Promote / Retest	Type of Test	Item Type	Calc. Use	Performance Levels (#)
OH	Ninth-Grade Proficiency Testing	CI, M, R, S, W	8> 9> 10> 11> 12	Y / N / Y	CRT, WP	MC, ER	N	2-Pass; Fail
OH	Twelfth-Grade Proficiency Testing	CI, M, R, S, W	12	N / N / N	CRT, WP	MC	Y	3 - Below Proficient; Proficient; Honors
OK	Criterion-Referenced Tests	M, R, S, W	5 > 6; 8 > 9; 11>12	N / N / Y	CRT, WP	MC, ER	N	2- Satisfactory; Unsatisfactory
SC	Basic Skills Assessment Program	M, R, S, W	3, 6, 8, 10, 11	Y / Y / N	CRT, WP	MC, ESW	N	2-Pass; Fail
TN	TCAP Competency Test	LA, M	9	Y / N / Y	CRT	MC	N	2-Pass; Fail
TX	TX Assessment of Academic Skills (TAAS)	M, R, S, SS, W	R&M-3-8, 10-12; W- 4, 8, 10 >11 >12; S&SS-8	Y / N / Y	CRT, WP	MC, ER	Y	3- Below Min. Expectations; Min. Expectations; Academic Recognition
UT	Core Assessment CRT Program	M, R, S	1-6, 7	N / N / N	CRT	MC	N	?
VA	Literacy Passport Test	M, R, W	6>7>8>9 >10 >11> 12	Y / Y / Y	CRT, WP	MC, FB/C, ESW	N	2-Pass; Fail

Subject Areas: LA = Language Arts; M = Mathematics; R = Reading; S = Science; SS = Social Studies; CI = Citizenship; CO = Communication; W = Writing.

Grade Level: A ">" sign indicates that students usually take the test in the first grade listed and those that do not pass apparently retake the exam in the subsequent grade(s) listed. A "-" sign indicates that all students in all grade levels between the two listed (e.g., 1-6) take the exams.

Required for HS Graduation, Promotion, and are Retests permitted: Y = Yes; N = No.

Type of Test: CRT = Criterion-Referenced Test; NRT = Norm-Referenced Test; WP = Writing Prompt; PA = Performance Assessment.

Item Type: MC = Multiple Choice, MC-SE = Multiple Choice with Student Explanation, FB/C = Fill-in the Blank or Cloze, SA = Short Answer, ER = Extended Response, ESW = Examples of Student Work.

Calculator Use (Was a calculator permitted on mathematics test?)**:** Y = Yes; N = No.

- "Failure" had to be one of the proficiency levels. This was established by examining the reported proficiency levels (CCSSO, 1996, pp. 205-207) or by inferring this from data indicating that all students in a state completed the test in a given grade level (e.g., grade 9) and then a smaller number of students completed the test in subsequent grades (CCSSO, 1996, pp. 119-130). Of the twenty-two testing programs with performance levels listed (CCSSO, 1996), 17 had only two levels named, while the remaining five indicated three proficiency levels. As expected, the proficiency levels of fifteen of the seventeen were simply called pass and fail. Those with three proficiency levels usually included two passing categories (e.g., honors, proficient) and a single "below minimum competency" category.

Twenty-four competency testing programs from nineteen states met these criteria.

Characteristics of Current Programs

Southern states are more heavily represented in the geographic distribution of these testing programs. Almost all of the southeastern states (Alabama, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee) except Kentucky and Arkansas included a minimum competency program in their state assessment system. Several states bordering on this southeast region—Virginia, Maryland, Texas, Oklahoma, Ohio—also employed a minimum competency program. The only northern states incorporating a minimum competency testing program in their state testing systems were New York and New Jersey.

Language arts and mathematics were the two most commonly tested content areas. However, science and/or social studies (including citizenship) were tested in eleven of the programs. The focus on mathematics and reading/language arts is likely related to fulfilling Title I accountability requirements. Further, reading and mathematics are often viewed as gateway subjects and policymakers tend to perceive the need to achieve minimum levels in reading and mathematics in order to be successful in other content areas.

Except for a few programs, most of the tests were geared toward late middle school and high school students. Approximately half of the programs allowed students to complete tests in subsequent years if they failed during the first administration. Tests targeted to these grade levels were intended to certify that students had the necessary skills to be ready for high school level content and to certify that high school graduates possess at least a minimum level of literacy and computational skills. The latter reason tends to be the most popular driving force among policymakers and the public for instituting competency testing programs (Reardon, 1996; Winfield, 1990).

Types of Tests and Items

Almost all of these states indicated they used criterion-referenced tests, although 16 of the 24 programs also included a writing prompt to assess students' writing. A few of the states indicated they used norm-referenced tests as part of the testing program. For this purpose—accountability of a certain competency level—there is very little difference between criterion (CRT) and norm referenced tests (NRT). Examination of NRTs and CRTs reveals that the items are essentially indistinguishable from one another. The only difference might be related to standard setting on the two types of tests. When norm-referenced tests are used for minimum competency purposes, students are often expected to score above a certain national percentile to be certified as competent. On the other hand, criterion referenced tests usually have an *a priori* level of performance that students are expected to reach if they are to be considered competent. However, we must emphasize that when both types of tests rely on multiple choice items, the two tests measure very similar levels of understanding.

Selection-type items (e.g., multiple-choice) or cloze procedures (e.g., fill-in-the-blank) can be used in certain circumstances to measure the application of meaningful content, but most multiple-choice items are geared toward testing content at the comprehension and knowledge (i.e., low cognitive) levels. On the other hand, open-ended items tend to be geared toward measuring application, analysis, synthesis, and evaluation skills—all considered higher level cognitive processes. When not professionally designed, these types of assessments can be used to only require students to construct a simple list of facts. In general though, multiple choice items tend to test lower-level cognitive processes than open-ended tasks. In the past, minimum competency testing programs were almost always comprised of selection-type items. As can be seen in Table 2, all of the competency testing programs listed are based largely on the use of multiple-choice or fill-in-the-blank items to assess all subject areas except writing.

Another check on the cognitive level required in these tests is to examine whether calculators are permitted on the mathematics portions of these tests. Current mathematics reform efforts call for increased use of calculators and other technology (NCTM, 1989). Typically, mathematics tests that permit the use of calculators test the application of mathematics concepts and skills. The converse is not necessarily true; mathematics tests that do not allow calculators can still test application and problem solving. However, assessments designed to measure basic computational skills would probably not permit calculator use. Eight of the twenty-four testing programs permitted a calculator to be used on the mathematics assessment. This means that calculators were not permitted on two-thirds of the tests.

Consequences Associated with Current State Minimum Competency Programs

The most consequential feature of these testing programs relates to the effect on those students who do not meet minimum expectations. Passing the competency test is required for promotion or is a graduation requirement in 20 of 24 testing programs. Students are able to retest in 19 of these 20 cases in order to meet the graduation requirement. In some states, (i.e., VA, MD, OH) students are given up to seven tries—starting as early as sixth grade—to pass the high school graduation test.

Even though 20 of these programs require students to pass test(s) in order to graduate, only three (Louisiana Criterion Referenced Test, South Carolina Basic Skills Assessment Program, and Virginia Literacy Passport Test) require students to pass exams for promotion to the next grade level. Two of these programs—Louisiana and South Carolina—require tests (and therefore the potential for grade retention) as early as third grade. Virginia's Literacy Passport Test begins with sixth grade students.

Summary of Current Programs

There is not a clear consensus among state programs about the purposes, consequences, and performance levels associated with minimum competency in these testing programs. The rhetoric describing these programs would have one believe that these tests no longer measure minimum levels of academic content; however, the evidence presented here contradicts much of the rhetoric. While it is admirable that two-thirds of these testing programs are now using direct writing prompts to assess students' writing achievement, essentially all of these programs rely on multiple choice tests to measure knowledge in the other subject areas tested. These tests are undoubtedly better than their predecessors in the 1970s, but many of them still appear to be a long way from assessing the types of knowledge and skills required for work in the 21st century. There is no question that many of these tests appear to be meeting an important political need; they enable policymakers to certify to their constituents that high school diplomas in their respective states now have meaning. We question whether meeting this political need is enough to justify the cost, time, and potential consequences associated with these testing programs.

Evaluations of Minimum Competency Testing Programs

Introduction and Evaluation Criteria

One of the problems with investigating the validity of minimum competency testing programs is that much of the rhetoric related to minimum competency testing programs is not based on any systematic analysis of empirical evidence. An often-cited problem with educational research is that one can seemingly always find a single study to support a particular perspective. However, the thoughtful reviewer will try to look fairly at the body of literature to synthesize an opinion supported by the preponderance of evidence. Both positive and negative findings from these studies are presented and we attempt to reach an evaluative conclusion about the efficacy of minimum competency testing.

When conducting an evaluation, it is important to establish criteria for successful or unsuccessful outcomes of the target of the evaluation. A validity study of minimum competency testing would have to examine both the positive and negative effects on curriculum and student learning. In addition to simply evaluating the positive and negative effects, an evaluator should determine whether these effects are intended or unintended. Shepard (1993) argued that simple correlations do not suffice for demonstrating positive effects; rather, there should be a theory of learning (and testing) that supports the use of minimum competency tests and the proposed remediation strategies. In other words, there needs to be some type of causal framework that leads to the expectation that the use of these large-scale minimum competency testing programs will lead to an increase in student learning. These criteria might appear to unfairly encumber proponents of a minimum competency testing program, but it could be argued that these expectations are just, considering the additional burden (e.g., time away from teaching and learning, stress, narrowing the curriculum) upon students and teachers (Shepard, 1993).

One of the major difficulties in evaluating a phenomenon such as minimum competency testing is that it is essentially impossible to conduct the ideal study. Most proponents of competency testing programs would not expect to observe program benefits immediately; the effects would likely require a long time-frame. Because students within schools are not randomly assigned to minimum competency testing programs, we need to judge the purported effects in light of other reasonable hypotheses. For example, if we are evaluating the influence of minimum competency testing programs on either achievement or dropping out of school during a five-year period, we need to recognize that other factors (e.g., increased national focus on education standards) could have produced the result. Good evaluations should try to account, either statistically or by establishing a comparison group, for other competing variables. This review was focused on these types of high quality studies.

Intended Effects

Increase in reading achievement

Several studies relied on data from the National Assessment of Educational Progress (NAEP) or the National Education Longitudinal Survey (NELS) to evaluate minimum competency testing programs. Linda Winfield's study of the relationship between reading achievement and minimum competency testing programs relied on the 1983-84 administration of NAEP (Winfield, 1990). Winfield focused on school-level effects and categorized schools as minimum competency testing or non-minimum competency testing schools based on responses to the principal questionnaire. The students from the two sets of schools were compared on reading achievement while statistically adjusting for gender, student age, region of the U.S., family background, student academic behaviors, and school-level socioeconomic status (SES). She found positive effects of minimum competency testing programs for students in grades eight and eleven, but no differences for fourth grade students. The differences, while statistically significant, were not very large in a practical sense (effect sizes were approximately 0.3 standard deviations, meaning that the average member of the higher scoring group is approximately at the 62nd percentile of the lower scoring groups). Winfield suggested that there were some systematic differences in her sample indicating that the minimum competency testing schools had a higher percentage of students in both gifted and remedial reading programs compared to schools without minimum competency testing programs. "This suggests that the identified 'MCT effect' may not be due solely to MCT but to other school-related conditions and characteristics" (Winfield, 1990, p. 168).

Increase in basic mathematical skills

Frederiksen (1994) used NAEP data to examine the effect of minimum competency testing programs on mathematics achievement. While Winfield conducted a school-level analysis, Frederiksen focused on the effects of state-mandated minimum competency testing programs. He analyzed trends in mathematics achievement between 1978 and 1986 using the same set of NAEP items. Frederiksen argued that 1978 was "before the minimum competency tests were in use, and the 1986 assessment occurred after minimum competency testing programs had been widely used—a before-and-after design." It is important to note that Frederiksen's assumptions are curious given the history of minimum competency testing programs discussed earlier in the paper. There is clear evidence that minimum competency testing programs had been in use prior to 1978. For example, Jaeger (1982) reported that at least 24 states had enacted minimum competency testing programs prior to 1978, 15 of which had been enacted prior to 1977.

Frederiksen classified states as high or low stakes depending upon the sanctions or rewards attached to their assessments, and classified mathematics problems as basic or higher-order.

However, "there were very few items that might be called higher-order skills, especially for the younger students" (Frederiksen, 1994, p. 8). He decided that "routine" and "nonroutine" would better describe the two groups of items. Mathematical skills that could be used automatically were considered routine. The few items requiring more thinking were called nonroutine (Frederiksen, 1994, p. 8). Frederiksen found that students from the 10 high-stakes states improved in their scores on routine problems compared with students from the 11 low-stakes states.

Unintended Effects

When conducting evaluations, it is not enough to simply examine whether the program does what its proponents say it is supposed to do. A thoughtful evaluator will look for and evaluate the unintended consequences of a given program or policy. These unintended consequences should carry as much weight in a validity investigation as the intended effects. This section presents an overview of some documented unintended consequences of minimum competency programs.

Lack of transfer to higher-order skills

Except for fourth graders, the advantage reported by Frederiksen on routine problems did not generalize to nonroutine problems. In Frederiksen's own words:

> *It seems reasonable to conclude that the use of minimum competency testing programs can have desirable influences on the performance of young students (i.e., 9 years old) as measured by NAEP—especially when high-stakes conditions prevail. It also seems reasonable to assume that for teenage students (i.e., 13 & 17 year olds) too much emphasis on teaching basic skills may indeed interfere with the teaching of higher-order thinking skills. If teaching basic skills interferes with acquiring nonroutine skills, they would surely interfere with teaching more advanced thinking abilities.* (Frederiksen, 1994, p. 14)

It must also be kept in mind that there were essentially no "higher-order" problems on the 1984 NAEP. NAEP did not include higher-order problems to a significant extent until the 1992 administration. The relatively weak transfer of basic skills to nonroutine items would be even weaker if transfer to truly higher-order problems was measured.

Increasing dropout rates

The most recent national evaluation of minimum competency testing programs used 1988 and 1990 data from the National Education Longitudinal Survey (NELS). Using a multiple regression

approach, Reardon (1996) statistically controlled for many factors thought to influence dropout rates (e.g., student age, socioeconomic status, grades, region in the U.S.) and then examined the relationship between the use of high-stakes minimum competency testing programs and dropping out. Typical of most complex issues, this relationship was not clear. In general, there appeared to be an increased risk for dropping out as a result of a minimum competency testing policy for students from low socioeconomic status (SES) schools. This relationship was reversed for higher SES schools. Some of this interaction could be due to sampling error because relatively few high SES schools have minimum competency testing programs. Perhaps, however, those high SES schools that use minimum competency tests are able to provide additional resources and remediation programs. Thus, the minimum competency testing programs might help identify students at risk for school failure and the school is able to target these students for remediation.

Reardon conducted a second set of analyses to try to get at the direction of causality—that is, does a minimum competency testing policy cause higher dropout rates or are schools with higher dropout rates more likely to institute minimum competency testing policies? He eliminated all students from this analysis who had been retained in eighth grade (based on his presumption that they failed the MCT) and then conducted the same regression analyses that he used earlier. He found no relationship between minimum competency testing policies and dropping out of school. This suggests that minimum competency testing, by causing students to be retained in eighth grade, tends to increase the rate of dropping out of school.

These analyses are a bit convoluted and in some sense speak to the problems of trying to conduct secondary analyses with large data sets. In both the NAEP and NELS studies described above, there were no variables to identify student performance on the minimum competency test and the researchers had to use some "back-door" approaches to answering there questions. Unfortunately, there are very few systematic studies that describe the specific relationship between student performance on the minimum competency test and other variables of interest.

There is one recent study using data from Florida's minimum competency graduation tests that might shed some light on this issue (Griffin & Heidorn, 1996). Data were analyzed from over 75,000 students, systematically sampled from Florida high schools during a five-year period (1986-91) in order to determine whether failing a minimum competency test influenced a student's decision to drop out of school. Overall, it appeared that failing a minimum competency test was related to an increased rate of dropping out of school. The most interesting aspect of this study is the interaction among minimum competency test scores, grade point average (GPA), and dropping out of school. Failing the minimum competency test had essentially no effect on dropout rates for students with low GPAs (i.e., <1.5), but curiously for those students with GPAs greater than 1.5, failing the minimum competency test was correlated with dropping out of school. It appears that many students with extremely low GPAs had already made the decision to drop out of school, so failing the minimum competency test had little additional effect. However

for students who were getting Cs and Bs, failing a minimum competency test seemed to influence their decision to leave school. Perhaps this failure created enough doubt in their academic ability and caused certain students to consider leaving school. This corroborates the results of Catterall's (1989) study where he found that students who failed a minimum competency test reported an increased likelihood of dropping out compared with those who passed the minimum competency test, while statistically controlling for other related variables.

Narrowing the curriculum

There have been several studies focusing on the effects of minimum competency testing on curriculum, teaching, and learning. Most of these studies have been critical of minimum competency tests because of their negative effects on curriculum and instruction. By definition, competency tests focus on low-level or basic skills-type learning. Because of this focus on basic skills, Shepard and Dougherty (1991) found that more than two-thirds of the teachers they interviewed emphasized basic skills more so than if there were no standardized tests. Lomax and colleagues (Lomax, West, Harmon, Viator, & Madaus, 1995) found similar sentiments echoed from teachers they interviewed. For example, a fifth grade teacher quoted by Lomax et al. said,

> *Testing restricts my teaching to a narrow range of objectives. I must follow precise curriculum objectives and make sure that each is covered. Eighty to ninety percent of time in math is based on preparation for proficiency. The exercises and tests are dull and unimaginative and do not encourage thinking or doing.* (Lomax et al., 1995, p. 179)

While the effects of minimum competency tests can be severe for white middle-class students, they are often disastrous for minority students. According to Lomax et al. (1995), this is primarily the result of an increased focus on test preparation, including the concomitant emphasis on basic skills instruction. Traditionally, minority students have not performed as well on standardized tests as their white peers and as a result have been tracked into classrooms with fewer resources and with lower quality teaching. Compounding the effects of their lower-track status, minority students are then subjected to the increased test preparation—more basic skills practice—in an effort to raise their scores. As a result, these students never get the opportunity to participate in the learning of cognitively complex process and concepts (Lomax et al., 1995).

> *These students rarely are given the opportunity to talk about what they know, to read real books, to write, or to construct and solve problems in mathematics, science, or other subjects. In short, they have been denied the opportunity to develop the capacities they will need for the future, in large part because commonly used tests are so firmly pointed at educational goals of the past.* (Darling-Hammond, 1994, p. 12)

Corruptibility of high stakes tests

Another serious problem with the use of high-stakes testing is that it can lead to unethical behavior by teachers, students, and administrators. Low scoring reliabilities reported from Kentucky are the most recent example. This is not a new phenomenon. Other researchers have reported that in order to meet the demands placed on them by policymakers and the public, school personnel and occasionally students, participated in unethical testing and test preparation activities, ranging from developing curriculum based on the content of the test to allowing students to "practice" on actual items from the test to be given (Haladyna, Nolen, & Haas, 1991). In the Wyoming context, if students were to be considered for grade retention on the basis of test results, and teachers and administrators were held accountable for the percentage of students passing the minimum competency exams, the possibility for corruption could be considerable.

Time away from teaching

Finally, in high stakes contexts, minimum competency tests may actually work directly against their intended effects. Proponents of these tests obviously want to see an increase in student achievement, but the time devoted to testing and test preparation actually reduces time and opportunities for learning. For example, in addition to approximately 4-5 days of testing, some teachers reported spending as much as 4 weeks each year on test preparation activities, taking valuable time away from instruction (Haladyna, Nolen, & Haas, 1991; Lomax et al., 1992; Shepard & Dougherty, 1991).

Practical Issues

Financial commitment

There are many practical issues that need to be addressed when discussing the implementation of a new testing program. The most obvious concern is the standard up-front costs of testing, including the cost of remedial programs. Exact costs of a minimum competency testing program would require more specific information from testing companies, but we can derive reasonable estimates based on other testing programs. The legislation targets students at three grade levels, which translates into approximately 25,000 school children in Wyoming. A standard, commercially available testing package to assess students in three subject areas would cost approximately $6-$10 per child. The higher cost would be associated with the use of a direct-writing prompt, while the lower estimate would use a survey battery of multiple-choice tests for all subject areas. For the sake of this exercise, an estimate of $8 per child is used, which translates into a cost to the state of Wyoming of $200,000 per year.

Estimating the cost of remediation is somewhat trickier because it depends on the specific program instituted. If students failing the minimum competency test are designated for grade retention, the cost would be exorbitant, but because it would not appear as a "line-item" much of this cost remains hidden. However, someone—most likely the state—would have to cover these costs. If a failure rate of 2.5% is assumed, similar to other minimum competency testing programs, 625 students would be expected to be retained each year. At a minimum these students would be in school an extra year. The state would be expected to reimburse the local district the per-pupil expenditure. If an average per-pupil expenditure of $5,500/student/year is assumed, then the cost to Wyoming taxpayers of retaining 625 pupils each year would be over three million dollars ($3,437,500)! While the state would be directly responsible for approximately 60% of this amount (i.e., $2,062,500), Wyoming taxpayers in one form or another would have to foot the entire bill.

There is an extensive body of literature demonstrating that grade retention is essentially always associated with negative effects for students (cf. Shepard & Smith, 1989). The cost, as demonstrated above, would make considering such a program prohibitive, but it should also be considered prohibitive on the basis on doing what is best for students. Another option would focus on remedial efforts for students failing the minimum competency test. One of the problems with providing remediation is that it will be focused on relatively few children in each school district at each of the three grade levels. Minimally, a remedial effort would require hiring paraprofessionals to serve school districts around the state. Hiring at least 26.5 paraprofessional educators (one each for the five largest districts and a half time each for the other 43 districts) at an average cost of $20,000/year would require approximately $530,000 per year. Perhaps some money could be used more efficiently in certain districts by using existing programs (e.g., Title I) to help meet the remediation needs of students failing the minimum competency test, but if an MCT program is going to be instituted, then appropriate funding should be provided so that it does not draw resources away from existing programs.

When a state-level minimum competency testing program is instituted, funding allocations should include the cost of a Department of Education coordinator and at least a one-half time administrative assistant. The statewide coordinator would supervise the yearly implementation of the testing, coordinate the remediation efforts, and collect appropriate data to be used for evaluative purposes. In Wyoming, the costs for the state coordinator would be approximately $50,000 per year ($40,000 salary, 25% benefits) and the half-time administrative assistant would be $15,625 per year ($12,500 salary, plus $3,125 benefits).

The cost of minimum competency testing in Wyoming, without any remediation or grade retention, therefore would be approximately $265,625 per year. Hiring 26.5 paraprofessionals for the least costly remediation effort—and probably the most appropriate—would cost an

additional $530,000 per year, bringing the total cost of a Wyoming minimum competency testing program to $795,625 per year.

Two Assessment Systems

As demonstrated above, the cost for instituting a minimum competency testing program and appropriate remediation would cost well over a half-million dollars each year. There is not an unlimited pool of education funding in any state and drawing off these resources would undoubtedly affect the legislatively mandated standards-based state assessment system. Developing and implementing an effective state assessment system will require more resources than have currently been allocated. Two assessment systems competing for the same resources means that neither will be done well.

More important than the diversion of resources will be the policy confusion resulting from having two contradictory testing systems. The goal of the standards-based testing system is to promote the teaching and learning of cognitively complex academic standards, while the minimum competency test will be geared toward measuring basic skills. Most likely, only one set of goals would take precedence. Based on previous research and experiences, it can be speculated that the test with the highest stakes would ultimately have the biggest effect on teaching and curriculum. If grade retention for students or negative evaluations of teachers were consequences of poor performance on the minimum competency test, then those types of stakes would likely drive the curriculum in a way that would not support learning of high-level content standards.

Finally, the design of the standards-based assessment system will include setting performance standards in the most informative and valid way possible. These performance standards should establish goals that citizens and policymakers would eventually like all students to reach while providing meaningful information about the current achievement levels of students. This will include information about students all along the achievement continuum, so that the identification of students at risk for academic failure would be accommodated by the proposed standards-based testing system. Using the standards-based system to identify low-achieving students will fulfill the goals of a minimum competency test, but will do so in a much more meaningful and less distracting way. Relatively few students fail a minimum competency test—estimates from other states indicate that it is between 2 and 8 percent of the students—but experiences with standards-based assessment systems suggests that many more students are performing below an acceptable level. This suggests that reform of teaching and learning should occur at the classroom and school level and not simply focus on those few students failing a minimum competency test. These low-performing students will undoubtedly benefit if the entire system was focused on bringing all students up to a high standard rather than trying to ensure that a few students can simply demonstrate minimum competency.

Discussion

Table 3 summarizes the findings of this document. The two-by-two table is organized according to the intended and unintended consequences, and by whether these consequences are theoretical or have been empirically documented. The results indicate that the potential (theoretical) benefits and the problems associated with minimum competency programs are fairly evenly distributed. The empirical findings, nevertheless, show that the negative consequences far outweigh the few positive results. For example, several studies (e.g., Frederiksen, 1994, Winfield, 1990) reported that student achievement improves when minimum competency testing programs are implemented. However, the measure of achievement in these cases is usually the same type of items as found on minimum competency tests. If the curriculum is focused on teaching basic skills, it is not surprising that students improve these skills, but we question the importance of this finding. The more important claim would address whether focusing on basic skills leads to an improvement in the types of higher-level knowledge and skills.

There have been no studies demonstrating that the use of minimum competency testing positively influences the types of skills and knowledge needed for work in the 21st century. There is an important reason why this type of information has not been reported. The use of minimum competency tests rests on an obsolete theory of learning; therefore we should not logically expect minimum competency programs to lead to learning higher-order thinking skills. Most learning theorists agree that sequential views of human learning where students are expected to progress along the same linear continuum until they reach subject-matter mastery do not hold anymore. Cognitive theories of learning suggest that humans interact with content to construct meanings and develop mental models. These models should not necessarily be measured in terms of "how much," rather the measurement should focus on "which one." Examination of these models suggests that people can learn and acquire concepts in very different ways. More importantly, up-to-date theories of learning indicate that students do not progress in lock step from simple skills to deep understanding. We do not dispute the importance of basic skills, but we disagree that they must always be acquired prior to learning more interesting content. Cognitive learning theories suggest that students can often best learn basic skills in the context of more complex problems. Until proponents of minimum competency tests can rationalize the expected causal relationship between these tests and student learning within a scientifically sound theoretical framework, we should not accept the claims regarding the supposed benefits of these tests.

Even if minimum competency tests do not lead to an increase of higher-order knowledge and skills, some might argue that if they are not causing harm, they should be used in hopes of potential benefit. This logic is flawed because proponents should have to provide evidence of positive consequences. The burden should not be on those opposing minimum competency

Table 3. Evaluation Matrix for Minimum Competency Testing Program

	Intended Effects	Unintended Consequences
Theoretical results	1. Ensure that all students have basic skills. 2. Give more meaning to high school diploma. 3. Curtail social promotion. 4. Relatively inexpensive policy. 5. Students move on to learn higher-order skills.	1. Increase dropout rate. 2. Increase in tracking. 3. Narrowing of the curriculum. 4. Exclusion of the teaching of higher-order thinking skills.
Empirical results	1. Improvement in basic mathematics skills (Frederiksen, 1994). 2. Increase in basic reading achievement (Winfield, 1990). 3. Increase in test scores (Frederiksen, 1994).	1. Increasing dropout rate for more successful students (Griffin & Heidorn, 1996, Reardon, 1996). 2. Narrowing curriculum (Lomax, West, Harmon, Viator, & Madaus, 1995; Shepard & Dougherty, 1991). 3. No gain or loss of higher-order thinking skills (Frederiksen, 1994). 4. Corruptibility of high stakes tests (Haladyna, Nolen, & Haas, 1991). 5. Time away from teaching and instruction (Haladyna, Nolen, & Haas, 1991; Lomax et al., 1992; Shepard & Dougherty, 1991). 6. Lack of transfer to high-order skills (Frederiksen, 1994). 7. Widening gap between educational "haves" and "have-nots" (Darling-Hammond, 1994; Lomax et al., 1992).

tests to offer evidence about the negative consequences of minimum competency testing programs.

Nevertheless, the studies reviewed earlier provide clear evidence of negative effects of minimum competency testing programs. For example, Griffin and Heidorn (1996) reported that minimum competency tests do not help those they are most intended to help, students at the lowest end of the achievement distribution. To compound the problem, these testing programs have had a negative effect on students with some promise. In addition to leading to an increase in dropout rates, one of the most deleterious effects of minimum competency tests is their influence on curriculum at the school and classroom level. This negative effect was recognized by one of the leading proponents of competency testing, James Popham:

> *When a state adopted a criterion-referenced test in the early days of competency testing, one of the distinguishing features of that test was that it clearly defined the nature of the skill being assessed. The strength of such a test was that it provided teachers with a clear depiction of what it was that they should be pursuing instructionally in their classes. But, as in most aspects of life, one's strength is almost always one's weakness. What happened over the years was that the very clarity that abetted instructional-design decisions also led to a kind of fragmented skills-focused approach that, in the end, may have been harmful to a number of students. Because the early competency tests promoted a clear understanding of what was to be taught, many teachers became preoccupied with a 'skill and drill' approach to teaching.* (Popham, 1991, p. 24)

While competency tests have improved in their most recent incarnation, they cannot help but have a narrowing effect on the curriculum, especially if the stakes are high enough. High stakes, such as grade retention or high school graduation, tend to cause teachers and schools to direct attention and other resources toward the test. Some argue that this is appropriate, especially if the test is measuring important learning goals (e.g., Resnick & Resnick, 1992). This is one of the major premises behind standards-based reform. The major differences between the type of measurement-driven instruction advocated by the Resnicks (and others) and minimum competency tests are the nature of the learning goals and the quality of the assessments. Years of evidence tell us that high stakes will drive curriculum and instruction, whether we like it or not. Many, and we are among them, believe that curriculum should not be driven in this basic skills direction.

One of the most fundamental problems with implementing a minimum competency and standards-based assessment systems at the same time is the resulting instructional confusion. Competency testing essentially contradicts current mandates for having students learn rigorous content standards. With dual systems and two different sets of expectations, teachers and students will be unsure of where they should aim. History informs us that the testing program with the

higher stakes will exert more of a driving force on curriculum and instruction. We doubt that any policymaker would like to see this confusion as a result of their well-intentioned efforts to improve an educational system.

On the other hand, one could argue that a proposed competency testing program will be designed to avoid this confusion and will be aligned with the state content standards, just geared toward measuring a lower performance level than the standards-based assessments. There are both philosophical and practical arguments to this position. Philosophically, it establishes a *de facto* two-track educational system. Those relegated to the minimum competency levels will be forever tracked into classes focused on basic skills while their peers in standards-based classes will be learning more demanding content. Thus, the gap between the high and low achievers will continue to widen.

Assessment plays a critical part within the context of current educational reforms. There are many benefits to be derived from setting standards for what students are to know and be able to do, and from assessing their performance against these standards. It is within this positive context of high expectations that states are most likely to move forward in their educational reform agenda.

References

Airasian, P. W. (1988). Measurement driven instruction: A closer look. *Educational Measurement: Issues and Practice, 7* (4), 6-11.

Baker, E. L. & Stites, R. (1990). Trends in testing in the USA. *Journal of Educational Policy, 5,* 139-157.

Bracey, G. W. (1987). Measurement-driven instruction: Catchy phrase, dangerous practice. *Phi Delta Kappan, 71,* 683-686.

Catterall, J. S. (1989). Standards and school dropouts: A national study of tests required for high school graduation. *American Journal of Education, 98,* 1-34.

Council of Chief State School Officers (CCSSO) (1997). *Annual survey of state student assessment programs, Fall 1996.* Washington, DC: Author.

Darling-Hammond, L. (1994). Performance-based assessment and educational equity. *Harvard Educational Review, 64,* 5-30.

Dewey, J. (1938/1963). *Experience and Education.* New York: Collier Books.

Frederiksen, N. (1994). *The influence of minimum competency tests on teaching and learning.* Princeton, NJ.: Educational Testing Service, Policy Information Center.

Glaser, R. & Silver, E. (1994). Assessment, testing, and instruction: Retrospect and prospect. *Review of Research in Education, 20,* 393-419.

Griffin, B. W. & Heidorn, M. H. (1996). An examination of the relationship between minimum competency test performance and dropping out of high school. *Educational Evaluation and Policy Analysis, 18,* 243-52.

Haladyna, T. M., Nolen, S. B., & Haas, N. S. (1991). Raising standardized achievement test scores and the origins of test score pollution. *Educational Researcher, 20* (5), 2-7.

Haney, W. & Madaus, G. (1989). Searching for alternatives to standardized tests: Whys, whats, and whithers. *Phi Delta Kappan, 70* (9), 683-87.

Jaeger, R. M. (1982). The final hurdle: Minimum competency achievement testing. In G. R. Austin & H. Garber (Eds.), *The rise and fall of national test scores* (pp. 223-246). New York: Academic Press.

Lomax, R. G., West, M. M., Harmon, M. C., Viator, K. A., & Madaus, G. F. (1995). The impact of mandated standardized testing on minority students. *Journal of Negro Education, 64*, 171-85.

Mislevy, R. J. (1996). *Some recent developments in assessing student learning*. Princeton, NJ: Educational Testing Service, Center for Performance Assessment.

National Commission on Excellence in Education. (1983). *A nation at risk : The imperative for educational reform: A report to the nation and the Secretary of Education, United States Department of Education*. Washington, DC : Author [Supt. of Docs., U.S. G.P.O. distributor].

National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

Popham, W. J. (1987). The merits of measurement-driven instruction. *Phi Delta Kappan, 68*, 679-682.

Popham, W. J. (1991). Interview on assessment with James Popham. *Educational Researcher, 20* (2), 24-27.

Reardon, S. F. (1996, April). *Eighth grade minimum competency testing and early high school dropout patterns*. Paper presented at the Annual Meeting of the American Educational Research Association, New York.

Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O'Conner (Eds.), *Changing assessments: Alternative views of aptitude, achievement and instruction.* Boston: Kluwer Academic.

Sarason, S. (1989). *The predictable failure of educational reform*. New York: Jossey Bass.

Shepard, L. A. (1988). *Should instruction be measurement-driven: A debate*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

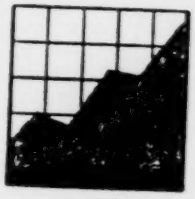
Shepard, L. A. (1991). Psychometricians' beliefs about learning. *Educational Researcher, 4*, 2-16.

Shepard, L. A. & Dougherty, K. C. (1991, April). *Effects of high-stakes testing on instruction.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Schimmel, J. (1997). The effects of increased graduation requirements on student achievement. Unpublished manuscript. University of Colorado, Boulder.

Tuman, M. C. (1979). Matthew Arnold and minimal competency testing. *The Journal of General Education, 31 (2)*, 122-28.

Winfield, L. F. (1990). School competency testing reforms and student achievement: Exploring a national perspective. *Educational Evaluation and Policy Analysis, 12*, 157-73.



NATIONAL
CENTER ON
EDUCATIONAL
OUTCOMES

University of Minnesota • 350 Elliot Hall
75 East River Road • Minneapolis, MN 55455
612.626.1530 • Fax 612.624.0879
http://www.coled.umn.edu/NCEO

The College of Education
& Human Development

UNIVERSITY OF MINNESOTA